

# Geometrical detection of weak non-Gaussianity upon coarse-graining

T.H. Beuman,<sup>1</sup> Ari M. Turner,<sup>2</sup> and V. Vitelli<sup>1,\*</sup>

<sup>1</sup>*Instituut-Lorentz for Theoretical Physics, Leiden University, NL 2333 CA Leiden, The Netherlands*

<sup>2</sup>*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA*

Measures of the non-Gaussianity of a random field depend on how accurately one is able to measure the field. If a signal measured at a certain point is to be averaged with its surroundings, or coarse-grained, the magnitude of its non-Gaussian component can vary. In this article, we investigate the variation of the “apparent” non-Gaussianity, as a function of the coarse-graining length, when we measure non-Gaussianity using the statistics of extrema in the field. We derive how the relative difference between maxima and minima – which is a geometrical measure of the field’s non-Gaussianity – behaves as the field is coarse-grained over increasingly larger length scales. Measuring this function can give extra information about the non-Gaussian statistics and facilitate its detection.

Keywords: non-Gaussian fields, random surfaces, extrema statistics, coarse-graining

## I. INTRODUCTION

As a consequence of the central limit theorem or of linear physics, many random fields are – at least as a first-order approximation – almost Gaussian. Examples of such fields can be found in various disciplines. The cosmic microwave background [1] is widely studied, but examples can also be found in optical speckle fields [2] and maps of brain activity [3]. Gaussian fields have universal properties that are well studied, including the statistics of its extrema and other singular points such as umbilics [4–8].

Deviations from Gaussianity are indicative of underlying nonlinear processes. Distilling these non-Gaussianities from the main signal can thus shed light on these nonlinear mechanisms. A standard way of testing and quantifying non-Gaussianity is to determine correlation functions beyond the second order. For Gaussian fields, these should, by Wick’s theorem, be factorizable into two-point correlation functions. A mismatch in these relations is therefore a tell-tale sign of non-Gaussianity.

However, this method requires detailed measurements of the field in question, and may therefore be impractical in some experimental settings. In [9–11], an alternative method was introduced, based on looking at the statistics of extrema and umbilical points, which in some circumstances can be less dependent on the accuracy with which the field can be probed. Removing one of these geometrical singularities or topological defects typically involves a nonlocal operation on the field that is unlikely to arise from local sources of noise or other perturbations.

In the present work, we investigate the effects that coarse-graining has on the statistics of extrema of a non-Gaussian field. This coarse-graining can represent the effect of imprecise measurements or averaging, but it can sometimes be helpful for the geometrical detection of non-Gaussianities. As illustrated in [9], a local non-Gaussian perturbation actually has an insignificant effect on the statistics of maxima and minima of the field if the resolution of the measurement is perfect. As a result, the perturbation would thus go undetected. As will be shown here, coarse-graining such a perturbed field can give rise to a sizable imbalance, thereby bringing the non-Gaussianity to light. For nonlocal perturbations, there is an imbalance at any resolution, but by comparing the number of maxima and minima at different scales one could get more information about the nonlocal correlations.

An example in which coarse-graining plays a different but intriguing role is gravitational lensing tomography. Images from distant galaxies are sheared due to mass present between that galaxy and us – this phenomenon is called weak gravitational lensing. Measuring the shear offers a window on the distribution of mass in the universe [12] – the singularities of the shear field for instance correspond to the umbilics of the projected gravitational potential [13]. It is a two-dimensional window though, providing only information about the projected gravitational potential. If one however incorporates the redshift – which can be translated to distance – a three-dimensional picture can be constructed [14]. The entire range of redshifts is divided into bins that can be processed individually. Therefore each bin yields an average of sources from a range of distances, rather than one specific distance. This can be interpreted as an effective source of coarse-graining.

---

\*Electronic address: vitelli@lorentz.leidenuniv.nl

The outline of this paper is as follows. The basics of Gaussian and non-Gaussian fields are given in section II. Section III sets up the mathematical framework for coarse-graining and discusses its effects. In section IV we derive what the imbalance between maxima and minima looks like in the limit that one coarse-grains over a large scale. This is demonstrated in section V, using an example case that allows the imbalance to be determined for arbitrary coarse-grain length scales. In section VI, the analytical result is compared to numerical simulations. Finally, section VII summarizes our findings.

## II. GAUSSIAN AND NON-GAUSSIAN FIELDS

### A. Gaussian fields

A homogeneous and isotropic Gaussian field is defined in terms of its Fourier components as

$$H(\vec{r}) = \sum_{\vec{k}} A(k) \cos(\vec{k} \cdot \vec{r} + \phi_{\vec{k}}). \quad (1)$$

The phases  $\phi_{\vec{k}}$  are independent random variables, uniformly distributed between 0 and  $2\pi$ . The *amplitude spectrum*  $A(k)$  depends only on the magnitude of the wave vector  $\vec{k}$ , not its angle – this ensures isotropy. There are no further constraints on the spectrum. Fields with different spectra can have different characteristics, but all still classify as Gaussian, which gives them some universal properties. For convenience, we will consider  $H$  to be normalized, such that

$$\langle H^2 \rangle = \langle H(\vec{r})^2 \rangle = \left( \prod_{\vec{k}} \int \frac{d\phi_{\vec{k}}}{2\pi} \right) H(\vec{r})^2 = 1. \quad (2)$$

A derivative of a Gaussian field is itself a Gaussian field, as the following example demonstrates:

$$\frac{\partial}{\partial x} H(\vec{r}) = \sum_{\vec{k}} A(k) k_x \cos(\vec{k} \cdot \vec{r} + \phi_{\vec{k}} - \frac{\pi}{2}). \quad (3)$$

The factor  $k_x$  can be absorbed into the spectrum and the  $\pi/2$  into the phase, so that the result still conforms to Eq. (1), although the field would no longer be isotropic.

Correlations between Gaussian fields can be completely characterized by the two-point correlations – higher-order correlations can be factorized into second-order ones as per Wick's theorem.

Equivalent to correlations are the so-called cumulants. The third-order cumulant of three variables for example can be expressed as

$$C_3(X_1, X_2, X_3) = \langle X_1 X_2 X_3 \rangle - \langle X_1 \rangle \langle X_2 X_3 \rangle - \langle X_2 \rangle \langle X_3 X_1 \rangle - \langle X_3 \rangle \langle X_1 X_2 \rangle + 2 \langle X_1 \rangle \langle X_2 \rangle \langle X_3 \rangle. \quad (4)$$

Gaussian variables have the characteristic property that all cumulants beyond the second order are zero.

### B. Non-Gaussian fields

In what follows, an uppercase  $H$  is used to designate a Gaussian field and a lowercase  $h$  for a perturbed Gaussian field. However,  $h$  is always taken to still be homogeneous and isotropic.

There are various types of perturbation. One example is  $h(\vec{r}) = H(\vec{r}) + f(H(\vec{r}))$ , where  $f$  is a (small-valued) function that depends on the original value of  $H$  at  $\vec{r}$  only. For this reason, this is called a *local* perturbation [9]. When the perturbation depends for instance on the gradient of  $H$ , the value of  $h(\vec{r})$  also encodes information about the surroundings of  $\vec{r}$ , which is why this type of perturbation is called *nonlocal*.

## III. COARSE-GRAINING

In general, mathematically, coarse-graining a field  $h(\vec{r})$  can be expressed as

$$\tilde{h}(\vec{r}) = \int d^2 \vec{u} K(\vec{u}) h(\vec{r} + \vec{u}), \quad (5)$$

with  $\int d^2\vec{u} K(\vec{u}) = 1$ .

When applied to a Gaussian field  $H(\vec{r})$ , the result is (in complex notation):

$$\begin{aligned}\tilde{H}(\vec{r}) &= \int d^2\vec{u} K(\vec{u}) \sum_{\vec{k}} A(k) e^{i(\vec{k} \cdot (\vec{r} + \vec{u}) + \phi_{\vec{k}})} \\ &= \sum_{\vec{k}} A(k) \left( \int d^2\vec{u} K(\vec{u}) e^{i\vec{k} \cdot \vec{u}} \right) e^{i(\vec{k} \cdot \vec{r} + \phi_{\vec{k}})}\end{aligned}\quad (6)$$

It is thus easily seen that coarse-graining only affects the amplitude spectrum of  $H$ , but not its Gaussianity.

Consider now what happens when applied to a non-Gaussian field  $h$ . Let us set  $K(\vec{r}) = f(r/l)/l^2$ , where  $l$  is a length scale that controls the size of the coarse-graining, while  $f(x)$  is a dimensionless function that goes to zero for  $x \gg 1$ . Let  $\xi$  be (a measure of) the correlation length of  $h$ , so that  $\langle h(\vec{r})h(\vec{r} + \vec{R}) \rangle$  vanishes when  $|\vec{R}| \gg \xi$ . If  $l \gg \xi$ , the coarse-graining effectively entails taking the average over a large number (of the order of  $(l/\xi)^2$ ) of independent regions, causing  $\tilde{h}$  to acquire Gaussian characteristics on account of the central limit theorem.

There is a link between coarse-graining and the deterministic Kardar-Parisi-Zhang (KPZ) equation that was investigated in [10]. This equation is a diffusion equation that reads

$$\frac{\partial h}{\partial t} = \nu \nabla^2 h + \frac{\lambda}{2} (\nabla h)^2. \quad (7)$$

Following the substitution  $u = \exp((\lambda/2\nu)h)$ , this transforms into  $\frac{\partial u}{\partial t} = \nu \nabla^2 u$ , with the solution

$$u(\vec{r}, t) = \int d^2\vec{r}' \frac{1}{4\pi\nu t} e^{-\frac{(\vec{r} - \vec{r}')^2}{4\nu t}} u(\vec{r}', 0). \quad (8)$$

This relation has the same structure as that of the formula for coarse-graining:  $u(\vec{r}, 0)$  can be identified as the original field and  $l = \sqrt{\nu t}$  as the coarse-graining scale. The dimensionless coarse-graining function is therefore

$$f(\rho) = \frac{1}{4\pi} e^{-\rho^2/4}. \quad (9)$$

In summary, coarse-graining a non-Gaussian field over a large scale causes it to obtain Gaussian characteristics. A Gaussian field remains Gaussian, regardless of the scale over which it is coarse-grained.

#### IV. LARGE SCALE LIMIT

In this section, the consequences of coarse-graining a non-Gaussian field  $h$  are investigated, in the limit that the coarse-graining scale becomes very large, as compared to the typical correlation length of the field. In particular, focus is put on the consequences for the densities of maxima and minima.

For a Gaussian field, the densities of maxima and minima are the same due to symmetry. For a non-Gaussian field, this is in general not the case. In [10, 11], a general expression was derived for the relative difference between the two densities for a perturbed Gaussian field, up to first order in the perturbation,

$$\Delta n \equiv \frac{n_{max} - n_{min}}{n_{max} + n_{min}} = \sqrt{\frac{6}{\pi\alpha}} \left( \frac{4}{3} \frac{\beta}{\sigma} + \frac{4}{9} \frac{\gamma}{\alpha} - \frac{10}{27} \frac{\delta}{\alpha} \right). \quad (10)$$

The parameters are second- and third-order correlations,

$$\sigma = \langle h_z h_{z^*} \rangle = C_2(h_z, h_{z^*}), \quad (11a)$$

$$\alpha = \langle h_{zz}^2 \rangle = C_2(h_{zz}, h_{zz}), \quad (11b)$$

$$\beta = \langle h_z h_{z^*} h_{zz} \rangle = C_3(h_z, h_{z^*}, h_{zz}), \quad (11c)$$

$$\gamma = \langle h_{zz}^3 \rangle = C_3(h_{zz}, h_{zz}, h_{zz}), \quad (11d)$$

$$\delta = \langle h_{zz} h_{zz^*} h_{zz} \rangle = C_3(h_{zz}, h_{zz^*}, h_{zz}). \quad (11e)$$

Here the subscripts denote partial differentiation, with  $\partial_z = \frac{1}{2}(\partial_x - i\partial_y)$  and  $\partial_{z^*} = \frac{1}{2}(\partial_x + i\partial_y)$ . In each cumulant, the variables are all taken at the same point  $\vec{r}$ , e.g.  $\sigma = C_2(h_z(\vec{r})h_{z^*}(\vec{r}))$ . Due to homogeneity, the choice of  $\vec{r}$  is irrelevant.

Note that it is not generally true that correlations and cumulants are identical. It is true though in this special case of correlations / cumulants up to third order between derivatives of a homogeneous field, as shall be demonstrated with  $\gamma$  as an example. Expressing the cumulant in correlations gives

$$\gamma = C_3(h_{zz^*}, h_{zz^*}, h_{zz^*}) = \langle h_{zz^*}^3 \rangle - 3\langle h_{zz^*} \rangle \langle h_{zz^*}^2 \rangle + 2\langle h_{zz^*} \rangle^3. \quad (12)$$

Note that, with the exception of the first, all terms carry a factor  $\langle H_{zz^*} \rangle$ , which can be expressed as

$$\langle h_{zz^*} \rangle = \partial_{z_1} \partial_{z_1^*} \langle h(\vec{r}) \rangle. \quad (13)$$

Since we consider  $h$  to be homogeneous,  $h(\vec{r})$  is constant and thus  $\langle h_{zz^*} \rangle = 0$ . This trick applies not only to  $\gamma$ , but to all five correlations in Eq. (11).

For the coarse-grained field  $\tilde{h}$ , the cumulants can be calculated in the following way (using  $\beta$  as an example):

$$\beta = C_3(\tilde{h}_z, \tilde{h}_{z^*}, \tilde{h}_{zz^*}) = \partial_{z_1} \partial_{z_2^*} \partial_{z_3} \partial_{z_3^*} C_3(\tilde{h}(\vec{r}_1) \tilde{h}(\vec{r}_2) \tilde{h}(\vec{r}_3)) \Big|_{\vec{r}_1 = \vec{r}_2 = \vec{r}_3} \quad (14)$$

The main ingredients that allow the cumulants to be calculated are thus  $C_2(\tilde{h}(\vec{r}_1), \tilde{h}(\vec{r}_2))$  (for  $\sigma$  and  $\alpha$ ) and the third-order equivalent (for  $\beta$ ,  $\gamma$  and  $\delta$ ).

Let  $\xi$  be a measure of the correlation length of  $h$ , in the sense that  $h(\vec{r}_1)$  and  $h(\vec{r}_2)$  can be said to be roughly uncorrelated when  $|\vec{r}_1 - \vec{r}_2| > \xi$ . If  $l \gg \xi$ , then  $\tilde{h}$  is nearly Gaussian (on account of the central limit theorem), so one can find the imbalance of maxima and minima using the approximation.

The second-order cumulant of  $\tilde{h}$  can be expressed as

$$C_2(\tilde{h}(\vec{r}_1), \tilde{h}(\vec{r}_2)) = \frac{1}{l^4} \iint d^2 \vec{R}_1 d^2 \vec{R}_2 C_2(h(\vec{R}_1), h(\vec{R}_2)) f\left(\frac{\vec{R}_1 - \vec{r}_1}{l}\right) f\left(\frac{\vec{R}_2 - \vec{r}_2}{l}\right). \quad (15)$$

Since  $C_2(h(\vec{R}_1), h(\vec{R}_2))$  is only appreciable when  $|\vec{R}_1 - \vec{R}_2| < \xi \ll l$ , the approximation  $f((\vec{R}_2 - \vec{r}_2)/l) \approx f((\vec{R}_1 - \vec{r}_2)/l)$  can be applied. This gives

$$C_2(\tilde{h}(\vec{r}_1), \tilde{h}(\vec{r}_2)) = \frac{1}{l^4} \iint d^2 \vec{R}_1 d^2 \vec{a} C_2(h(\vec{0}), h(\vec{a})) f\left(\frac{\vec{R}_1 - \vec{r}_1}{l}\right) f\left(\frac{\vec{R}_1 - \vec{r}_2}{l}\right), \quad (16)$$

where  $\vec{a} = \vec{R}_2 - \vec{R}_1$  and use was made of the homogeneity of  $h$ . This integration can now be split into two parts:

$$\begin{aligned} C_2(\tilde{h}(\vec{r}_1), \tilde{h}(\vec{r}_2)) &= \frac{1}{l^4} \int d^2 \vec{a} C_2(h(\vec{0}), h(\vec{a})) \int d^2 \vec{R}_1 f\left(\frac{\vec{R}_1 - \vec{r}_1}{l}\right) f\left(\frac{\vec{R}_1 - \vec{r}_2}{l}\right) \\ &= \frac{1}{l^2} I_2 K_2\left(\frac{\vec{r}_1}{l}, \frac{\vec{r}_2}{l}\right), \end{aligned} \quad (17)$$

where

$$I_2 \equiv \int d^2 \vec{a} C_2(h(\vec{0}), h(\vec{a})), \quad (18)$$

and

$$K_2(\vec{\rho}_1, \vec{\rho}_2) \equiv \int d^2 \vec{v} f(\vec{v} - \vec{\rho}_1) f(\vec{v} - \vec{\rho}_2). \quad (19)$$

For the third-order correlation an analogous derivation can be made. The result is

$$C_3(\tilde{h}(\vec{r}_1), \tilde{h}(\vec{r}_2), \tilde{h}(\vec{r}_3)) = \frac{1}{l^4} I_3 K_3\left(\frac{\vec{r}_1}{l}, \frac{\vec{r}_2}{l}, \frac{\vec{r}_3}{l}\right), \quad (20)$$

with

$$I_3 \equiv \iiint d^2 \vec{a} d^2 \vec{b} C_3(h(\vec{0}), h(\vec{a}), h(\vec{b})), \quad (21)$$

and

$$K_3(\vec{\rho}_1, \vec{\rho}_2, \vec{\rho}_3) \equiv \int d^2 \vec{v} f(\vec{v} - \vec{\rho}_1) f(\vec{v} - \vec{\rho}_2) f(\vec{v} - \vec{\rho}_3). \quad (22)$$

Note in particular that  $I_2$  and  $I_3$  depend on  $h$  only, whereas  $K_2$  and  $K_3$  depend only on  $f$ . Also note that none of these terms depends on  $l$ .

For the correlations, the  $z$ - and  $z^*$ -derivatives, as used in Eq. (14), act only on  $K_2(\frac{\vec{r}_1}{l}, \frac{\vec{r}_2}{l})$  and  $K_3(\frac{\vec{r}_1}{l}, \frac{\vec{r}_2}{l}, \frac{\vec{r}_3}{l})$ . Each derivative introduces a factor  $1/l$  as a result of the chain rule. It can therefore already be deduced how the relevant correlations scale with  $l$ :

$$\sigma \sim l^{-4}, \quad \alpha \sim l^{-6}, \quad \beta \sim l^{-8}, \quad \gamma, \delta \sim l^{-10}, \quad (23)$$

and therefore  $\Delta n \sim 1/l$ . More precisely, we have the following:

$$\lim_{l/\xi \rightarrow \infty} \Delta n = \frac{c_h c_f}{l}, \quad (24)$$

where  $c_h = I_3/I_2^{3/2}$  is a parameter that depends on the statistics of  $h$  only, and  $c_f$  is a parameter that depends on the coarse-graining function  $f$  only.

## V. EXAMPLE

### A. Large scale limit

As an example, let us consider the coarse-grain function from Eq. (9). This gives

$$K_2(\vec{\rho}_1, \vec{\rho}_2) = \frac{1}{8\pi} e^{-\frac{1}{8}(\vec{\rho}_1 - \vec{\rho}_2)^2}, \quad (25a)$$

$$K_3(\vec{\rho}_1, \vec{\rho}_2, \vec{\rho}_3) = \frac{1}{48\pi^2} e^{-\frac{1}{4}(\vec{\rho}_1^2 + \vec{\rho}_2^2 + \vec{\rho}_3^2 - \frac{1}{3}(\vec{\rho}_1 + \vec{\rho}_2 + \vec{\rho}_3)^2)}. \quad (25b)$$

Applying the method as exemplified in Eq. (14), we obtain

$$\beta = -\frac{I_3}{6912\pi l^8}. \quad (26)$$

Calculating all relevant cumulants ultimately leads to

$$\Delta n = \frac{2^{9/2}}{3^{11/2}\pi} \frac{I_3}{I_2^{3/2}} \frac{1}{l}, \quad (27)$$

for the KPZ-inspired Gaussian coarse-grain function Eq. (9).

To get the parameter  $c_h$ , we use a non-Gaussian field of the form  $h = H + \varepsilon H^2$ , where  $H$  is Gaussian field with a given two-point correlation function  $\langle H(\vec{r}_1)H(\vec{r}_2) \rangle$ , and  $\varepsilon$  is a small constant. The two-point correlation function of  $h$  differs from that of  $H$  only in second order of  $\varepsilon$  – this difference will be ignored.

The third-order cumulant of  $h$  is zero in leading order, since  $H$  is Gaussian, but in first order we find

$$\begin{aligned} C_3(h(\vec{r}_1), h(\vec{r}_2), h(\vec{r}_3)) &= C_3(\varepsilon H(\vec{r}_1)^2, H(\vec{r}_2), H(\vec{r}_3)) \\ &\quad + C_3(H(\vec{r}_1), \varepsilon H(\vec{r}_2)^2, H(\vec{r}_3)) \\ &\quad + C_3(H(\vec{r}_1), H(\vec{r}_2), \varepsilon H(\vec{r}_3)^2) \end{aligned} \quad (28)$$

This can be expanded with the help of Wick's theorem, e.g.:

$$\begin{aligned} C_3(H(\vec{r}_1)^2, H(\vec{r}_2), H(\vec{r}_3)) &= \langle H(\vec{r}_1)^2 H(\vec{r}_2) H(\vec{r}_3) \rangle - \langle H(\vec{r}_1)^2 \rangle \langle H(\vec{r}_2) H(\vec{r}_3) \rangle \\ &= 2 \langle H(\vec{r}_1) H(\vec{r}_2) \rangle \langle H(\vec{r}_1) H(\vec{r}_3) \rangle, \end{aligned} \quad (29)$$

Therefore, the result is

$$\begin{aligned} C_3(h(\vec{r}_1), h(\vec{r}_2), h(\vec{r}_3)) &= 2\varepsilon (\langle H(\vec{r}_1) H(\vec{r}_2) \rangle \langle H(\vec{r}_1) H(\vec{r}_3) \rangle + \langle H(\vec{r}_2) H(\vec{r}_3) \rangle \langle H(\vec{r}_2) H(\vec{r}_1) \rangle \\ &\quad + \langle H(\vec{r}_3) H(\vec{r}_1) \rangle \langle H(\vec{r}_3) H(\vec{r}_2) \rangle) \end{aligned} \quad (30)$$

If we consider the Gaussian two-point correlation function

$$\langle H(\vec{r}_1) H(\vec{r}_2) \rangle = e^{-\frac{1}{2} k_0^2 (\vec{r}_1 - \vec{r}_2)^2}, \quad (31)$$

– which corresponds to the spectrum  $A(k) \sim \exp(-k^2/(4k_0^2))$  – we get

$$I_2 = \int d^2 \vec{a} C_2(0, \vec{a}) = \int d^2 \vec{a} \langle H(0) H(\vec{a}) \rangle = \frac{2\pi}{k_0^2}, \quad (32)$$

and

$$\begin{aligned} I_3 &= \iint d^2 \vec{a} d^2 \vec{b} C_3(h(\vec{0}), h(\vec{a}), h(\vec{b})) \\ &= \iint d^2 \vec{a} d^2 \vec{b} 2\varepsilon \left( e^{-\frac{1}{2} k_0^2 (\vec{a}^2 + \vec{b}^2)} + e^{-\frac{1}{2} k_0^2 (\vec{a}^2 + (\vec{a} - \vec{b})^2)} + e^{-\frac{1}{2} k_0^2 (\vec{b}^2 + (\vec{a} - \vec{b})^2)} \right) \\ &= \frac{24\pi^2 \varepsilon}{k_0^4}. \end{aligned} \quad (33)$$

This gives  $c_h = I_3/I_2^{3/2} = 6\sqrt{2\pi}\varepsilon/k_0$ .

Combined with the coarse-grain function as given above, we thus get

$$\Delta n \rightarrow \frac{c_f c_h}{l} = \frac{64}{81\sqrt{3\pi}k_0} \frac{\varepsilon}{l}. \quad (34)$$

## B. Analytic result

The separation of the dependence on  $f$  and  $h$ , as displayed in Eq. (24), is only valid in the limit of  $l \gg \xi$ . In general however,  $f$  and  $h$  can no longer be treated separately. Only in very specific cases is it possible to calculate  $\Delta n$  for arbitrary  $l$ . Not coincidentally, the  $f$  and  $h$  chosen in the previous section allow precisely this.

For example, the exact expression for  $\beta$  is

$$\begin{aligned} \beta &= \langle \tilde{h}_z \tilde{h}_{z^*} \tilde{h}_{zz^*} \rangle \\ &= \partial_{z_1} \partial_{z_2^*} \partial_{z_3} \partial_{z_3^*} \iiint d^2 \vec{u}_1 d^2 \vec{u}_2 d^2 \vec{u}_3 K(\vec{u}_1) K(\vec{u}_2) K(\vec{u}_3) \\ &\quad \langle h(\vec{r}_1 + \vec{u}_1) h(\vec{r}_2 + \vec{u}_2) h(\vec{r}_3 + \vec{u}_3) \rangle \Big|_{\vec{r}_1 = \vec{r}_2 = \vec{r}_3}. \end{aligned} \quad (35)$$

The three-point correlation can be expanded in the same way as before (see Eq. (28) and Eq. (29)). The final result is

$$\beta = -\frac{k_0^4 \varepsilon}{2(1 + 2k_0^2 l^2)^2 (1 + 6k_0^2 l^2)^2}, \quad (36)$$

Determining and combining all the correlations gives the following result, which is exact with respect to  $l$  but still perturbative with respect to  $\varepsilon$ :

$$\Delta n = \frac{64a^3(1 + 4a)^{7/2}\varepsilon}{\sqrt{3\pi}(1 + 2a)^3(1 + 6a)^4} \quad (37)$$

where  $a \equiv k_0^2 l^2$ . In the limit of large  $l$  ( $a \gg 1$ ) we find that it matches the perturbative result.

One may also note that Eq. (37) matches the result from the deterministic KPZ equation for the Gaussian spectrum [10], following the substitution  $\nu t \rightarrow l^2$ .

## VI. NUMERICAL TESTS

### A. Setup

The validity of Eq. (37) was checked using computer simulations. The setup of the simulations and the identification of the extrema is identical to the process outlined in [9].

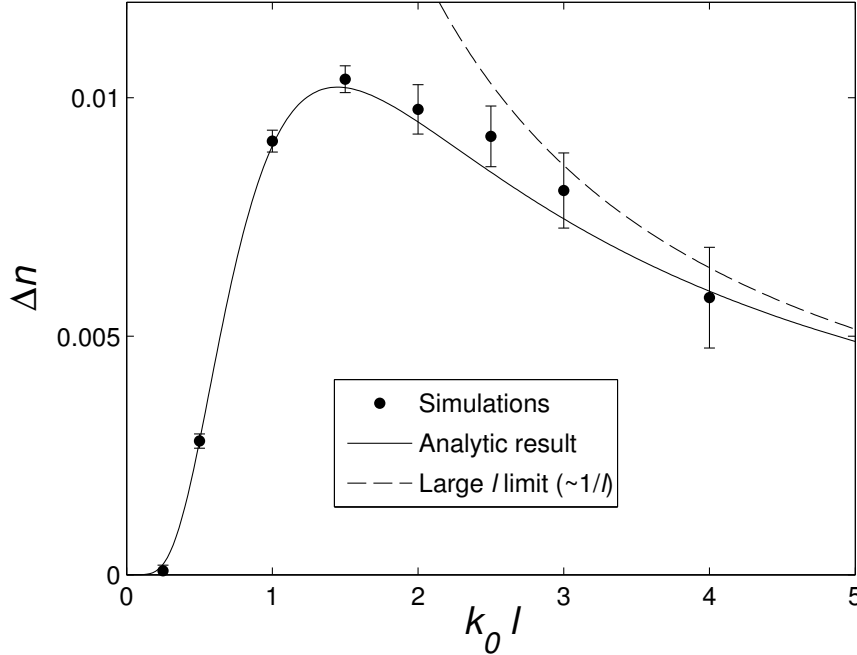


Figure 1: The imbalance between maxima and minima  $\Delta n$  for a field  $h = H + \varepsilon H^2$ , where  $H$  is a Gaussian field with a Gaussian power spectrum (with typical wavelength  $k_0^{-1}$ ), coarse-grained with a Gaussian function  $f(r) = \exp(r^2/(4l^2))$ . The solid line is the exact result Eq. (37) (perturbative in  $\varepsilon$  but not in  $l$ ), while the data points stem from simulations. The dashed line is the theoretical result for large  $l$ , Eq. (34).

For each data point – corresponding to a particular value of  $l$  – thousands of Gaussian fields  $H(\vec{r})$  were generated, following Eq. (1). In each case, hundreds of waves were summed, each with an amplitude in accordance with the desired spectrum and a random value for the phase. The values of  $H$  were evaluated for the points of a square grid of size  $L$ . Periodic boundary conditions were enforced to reduce finite size effects, which entail that the components of the selected wave vectors  $k$  were all multiples of  $\frac{2\pi}{L}$ .

To this field, the perturbation  $\varepsilon H^2$  was added, with  $\varepsilon = 0.1$ . This new field  $h$  was then coarse-grained, after which the extrema were identified. For this identification, the first and second derivatives of the coarse-grained field  $\tilde{h}$  were used, which were determined by calculating the derivatives of the original Gaussian field  $H$  and letting them undergo the equivalent process.

## B. Results

Figure 1 shows the theoretical result, as well as results from simulations at various values of  $l$ . As can be seen, there is an excellent agreement between the two. Also shown is the prediction of Eq. (34), illustrating the large  $l$  limit, which matches well for  $k_0 l \gg 1$ .

An interesting point is that, for no coarse-graining at all, the imbalance is very close to zero. This general feature of local perturbations of the type  $h = H + f(H)$  was already established in detail in [9]. Measuring the imbalance between maxima and minima thus does not reveal the non-Gaussianity of  $h$ . However, it is clear from figure 1 that coarse-graining may significantly increase the imbalance to measurable values, thereby not only granting the possibility of detecting non-Gaussianity, but also potentially identifying the size and type of the perturbation.

## C. Large scale coarse-graining

It is difficult to test the formula for a large coarse-graining (Eq. (24)) accurately in a numerical setting, since the size of the system is limited.

As said, the periodic boundary conditions are enforced by only using waves with wave vectors  $\vec{k}$  for which the components are multiples of  $\frac{2\pi}{L}$ , where  $L$  is the system size. As long as  $L$  is large, this quantization has a high enough resolution to be of no significant source of error.

Now consider what happens when it is coarse-grained. We already saw in Eq. (6) that, effectively, its amplitude spectrum changes:

$$\tilde{A}(k) = A(k) \left( \int d^2\vec{u} K(\vec{u}) e^{i\vec{k}\cdot\vec{u}} \right) = A(k) \left( \int d^2\vec{\rho} f(\rho) e^{i\vec{k}\cdot\vec{\rho}} \right). \quad (38)$$

In the limit that  $lk \gg 1$ , the phase factor causes the integral to vanish. Hence, for large  $l$ , only the waves with small wave vector  $k$  (in the order of  $1/l$  or less) prevail. This is the technical justification of the statement that coarse-graining causes a field to become smoother, and thus dominated by long waves.

However, in combination with the periodic boundary conditions, this means that – in the case that  $l$  becomes comparable to  $L$  – there are only a few wave vectors left that are of importance from the coarse-graining point of view. The accuracy with which the simulated coarse-grained field represents an actual field thus becomes compromised. Therefore,  $L$  should be larger than  $l$ . Increasing  $L$  however naturally increases computation time. As a result, probing large coarse-grain scales indirectly requires a lot of computation time, making it difficult to properly explore the regime in which the imbalance  $\Delta n$  decays as  $1/l$ .

## VII. CONCLUSIONS

Coarse-graining a non-Gaussian field has the effect of giving it Gaussian characteristics, as the coarse-graining scale goes to infinity. More precisely, when this scale  $l$  is significantly larger than the correlation length of the field, the imbalance between maxima and minima – which is zero for Gaussian fields – scales as  $1/l$ . The corresponding constant factor can be written as the product of two independent scalars: one depends on the field only, whereas the other depends on the coarse-graining function only.

Coarse-graining a signal on purpose can also be useful, because the imbalance between maxima and minima depends on the length scale of the coarse-graining for a non-Gaussian field. For example, locally perturbed fields, such as  $h = H + \varepsilon H^2$ , where  $H$  is Gaussian, do not show a significant imbalance between maxima and minima (if the resolution is perfect). However, coarse-graining – which would not produce an effect for Gaussian fields – creates an imbalance allowing  $\varepsilon$  to be measured. In general, coarse-graining by various amounts can give a multitude of data that can be used to shed light on some unknown parameters of the perturbation.

- 
- [1] S. Dodelson, *Modern Cosmology* (Academic Press, Amsterdam, 2003).
  - [2] F. Flossmann, K. O'Holleran, M. R. Dennis, and M. J. Padgett, *Phys. Rev. Lett.* **100**, 203902 (2008).
  - [3] K. J. Worsley, S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans, *Human Brain Mapping* **4**, 58 (1996).
  - [4] M. Longuet-Higgins, *Phil. Trans. R. Soc. Lond. A* **249**, 321 (1957).
  - [5] M. S. Longuet-Higgins, *Phil. Trans. R. Soc. Lond. A* **250**, 157 (1957).
  - [6] M. V. Berry and J. H. Hannay, *J. Phys. A: Math. Gen.* **10**, 1809 (1977).
  - [7] M. R. Dennis, *J. Phys. A: Math. Gen.* **36**, 6611 (2003).
  - [8] M. R. Dennis, *Optics Letters* **33**, 2572 (2008).
  - [9] T. H. Beuman, A. M. Turner, and V. Vitelli, *Phys. Rev. E* **88**, 012115 (2013).
  - [10] T. H. Beuman, A. M. Turner, and V. Vitelli, *Phys. Rev. E* **87**, 022142 (2013).
  - [11] T. H. Beuman, A. M. Turner, and V. Vitelli, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 19943 (2012).
  - [12] H. Hoekstra and B. Jain, *Ann. Rev. Nucl. Part. Sci.* **58**, 99 (2008).
  - [13] V. Vitelli, B. Jain, and R. D. Kamien, *JCAP* **09**, 034 (2009).
  - [14] W. Hu, *ApJL* **522**, L21 (1999).